

### **Remarks**

Claims 1-16 remain pending in this application after entry of this paper. In the Office Action dated September 24, 2003, the Examiner rejected claims 1 and 9 under 35 U.S.C. § 102(b) as being anticipated by Huang et al. (U.S. Patent No. 5,913,193). The Examiner also rejected claims 2-8 and 10-16 under 35 U.S.C. § 103(a) as being unpatentable over Huang et al. (U.S. Patent No. 5,913,193). The invention is believed to be patentable.

Claim 1 recites a method for converting text to concatenated voice by utilizing a digital voice library and a set of playback rules. The digital voice library includes a plurality of speech items and a corresponding plurality of voice recordings. Each speech item corresponds to at least one available voice recording. Multiple voice recordings correspond to a single speech item and represent various inflections of that single speech item. The method includes receiving text data, converting the text data into a sequence of speech items in accordance with the digital voice library. The method further comprises establishing multiple voice recordings in the digital voice library that corresponds to a single inflection of a single speech item, for a plurality of inflections of a plurality of speech items, that represent various ligatures for the single inflection of the single speech item with adjacent speech items.

This aspect of the invention is exemplified in the specification at page 31, line 13 through page 33, line 12, among other places.

Huang fails to anticipate the invention as set forth by Claim 1. Huang fails to describe "multiple voice recordings in the digital voice library that correspond to a single inflection of a single speech item, for a plurality of inflections of a plurality of speech items, that represent various ligatures for the single inflection of the single speech item with adjacent speech items." Huang does describe a method and system of runtime acoustic unit selection for speech synthesis. Huang discusses acoustic unit selection for speech synthesis and attempts to minimize the spectral distortion between boundaries of adjacent instances. Huang discusses a training phase where the highest probability instances representing diverse phonetic contexts

are chosen (Col. 2, ll. 65-67). Huang further describes that the naturalness of synthesized speech is improved by providing multiple instances of an acoustic unit, such as a diphone. (Col. 9, ll. 58-60.)

Although Huang describes providing multiple instances of an acoustical unit, with the diphone being an exemplary acoustical unit, Huang fails to discuss multiple voice recordings in the digital voice library that corresponds to a single inflection of a single speech item that represent various ligatures for the single inflection of the single speech item with adjacent speech items in the particular combination claimed by claim 1.

Huang only describes the creation of a robust library with multiple instances of an acoustical unit but fails to suggest the specific technique set forth by independent claim 1 involving inflections and ligatures.


For these reasons, claim 1 is believed to be patentable. The Examiner makes reference to diphones, but diphones are merely an acoustic unit and the use thereof fails to suggest the invention as set out in claim 1.

Claim 9 is an independent method claim and is believed to be patentable for the same reasons as given above for claim 1.

Claims 2-8 and 10-16 are dependent claims and are believed to be patentable for reasons given above with respect to the base claims.

Respectfully submitted,

**ELIOT M. CASE ET AL.**

By   
Jeremy J. Curcuri, Reg. No. 42,454  
Attorney for Applicants

Date: January 26, 2004

**BROOKS KUSHMAN P.C.**  
1000 Town Center, 22nd Floor  
Southfield, MI 48075-1238  
Phone: 248-358-4400; Fax: 248-358-3351